MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2017/2018

TIS3351 – ADVANCED DATABASE

(All sections / Groups)

13 MARCH 2018
9.00 a.m – 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This Question paper consists of 7 pages with 4 questions only.
2. Attempt **ALL** questions. All questions carry equal marks and the distribution of the marks for each question is given.
3. Please print all your answers in the Answer Booklet provided.

MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2017/2018

TIS3351 – ADVANCED DATABASE

(All sections / Groups)

13 MARCH 2018
9.00 a.m – 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This Question paper consists of 7 pages with 4 questions only.
2. Attempt **ALL** questions. All questions carry equal marks and the distribution of the marks for each question is given.
3. Please print all your answers in the Answer Booklet provided.

Question 1 [10 marks]

a) Functionalities of data warehouse include pivot, drill-down, roll-up, slice and dice, and selection. Explain drill-down and roll-up using an example each.

[2 marks]

b) *Golden Leaf* is a company selling health supplements within South East Asia (SEA). The company sources for the ingredients from vendors located at different parts of the world. It manufactures the products at its manufacturing plant located in Johor. The products are then sold to many distributors within SEA. The supplements can be grouped into different categories, such as vitamins, beauty, kids and senior citizens.

The company has been established for 20 years. In order to improve its business, the company plans to implement a data warehouse running on relational online analytical processing (ROLAP) system to garner business intelligence.

i) Based on the case study, identify the potential facts (measurable attributes) and dimensions for the data warehouse.

[2 marks]

ii) Draw a star or snowflake schema for the data warehouse based on the fact(s) and dimensions that you have identified. Include suitable attributes, primary keys and foreign key(s) in the fact and dimension tables.

[6 marks]

Question 2 [10 marks]

Relational OLAP (ROLAP) uses relational or extended-relational database management system (DBMS) to store and manage warehouse data and OLAP middle ware. Given the following relational schema of a ROLAP, answer questions (a) to (d). Primary keys are underlined while foreign keys are in *italic*.

LOCATION (Location_ID, Description)

BRANCH (Branch_ID, Loc_Detail, *Location_ID*)

SALESPERSON (SP_ID, First_Name, Last_Name, *Branch_ID*)

PRODUCT (Product_ID, Description, Unit_Price, Unit_Cost)

TIME (Time_ID, Month_Name, Year)

SALES (*SP_ID*, *Time_ID*, *Product_ID*, Sales_Quantity)

Continued...

The sample records in the tables:

LOCATION

Location_ID	Description
100	East Malaysia
200	West Malaysia
300	Thailand
400	Singapore

BRANCH

Branch_ID	Loc_Detail	Location_ID
B0001	Sandakan	100
B0002	Miri	100
B0003	Seremban	200
B0004	Kuala Lumpur	200
B0005	Bangkok	300
B0006	Hatyai	300
B0007	Penang	200
B0008	Woodlands	400

SALESPERSON

SP_ID	First_Name	Last_Name	Branch_ID
SP100	Andy	Chang	B0001
SP101	Mahmud	Omar	B0001
SP102	Rina	Khan	B0002
SP103	Saravanan	Raju	B0003
SP104	Henry	Choo	B0004
SP105	Jay	Kwan	B0005
SP106	Hisham	Mohamad	B0005
SP107	Praya	Kam	B0007
SP108	Wendy	Prasanam	B0006
SP109	Ian	Pang	B0008
SP110	Sienna	Woo	B0003

PRODUCT

Product_ID	Description	Unit_Price	Unit_Cost
10001	Learn English Level 1	200.00	150.00
10002	Learn English Level 2	200.00	150.00
10003	Learn Mandarin Level 1	250.00	160.00
10004	Learn Mandarin Level 2	250.00	160.00
10005	Learn Mandarin Level 3	280.00	180.00
10006	Learn Japanese Beginner	220.00	140.00
10007	Learn Korean Beginner	220.00	140.00
10008	Learn Mandarin Set	720.00	480.00

Continued...

TIME

Time_ID	Month_Name	Year
M001	January	2016
M016	April	2017
M018	June	2017
M022	December	2017

SALES

SP_ID	Time_ID	Product_ID	Sales_Quantity
SP100	M001	10001	2500
SP101	M016	10006	5000
SP106	M018	10008	1200
SP109	M001	10007	890
SP103	M018	10008	1800
SP103	M022	10008	1100
SP105	M016	10003	550
SP108	M016	10003	200

a) Write a SQL statement to produce the following output. The table lists the description of products and its total quantity sold.

Description	Total_Sold_Quantity
Learn English Level 1	2500
Learn Japanese Beginner	5000
Learn Korean Beginner	890
Learn Mandarin Level 1	750
Learn Mandarin Set	4100

[2 marks]

b) Write a SQL statement to list the first name and last name of salesperson who has sold 'Learning Japanese Beginner', as shown below.

Salesperson
Hisham Mohamad
Saravanan Raju

[3 marks]

Continued...

c) Write a SQL statement to produce the following result:

Description	Loc_Detail	Total_Quantity_Sold
-	-	13240
East Malaysia	-	7500
Singapore	-	890
Thailand	-	1950
West Malaysia	-	2900
East Malaysia	Sandakan	7500
Singapore	Woodlands	890
Thailand	Bangkok	1750
Thailand	Hatyai	200
West Malaysia	Seremban	2900

[3 marks]

d) Write a SQL statement to create a user defined function that returns the Branch ID and total quantity of products sold by each branch.

BranchID	Quantity_Sold
B0001	7500
B0003	2900
B0005	1750
B0006	200
B0008	890

[2 marks]

Question 3 [10 marks]

a) One of the concerns in distributed database design is data fragmentation. Given SALES table in Question 2, between horizontal and vertical fragmentation, which is more suitable? Explain why. Then, show how you perform the chosen fragmentation mode on SALES table.

[3 marks]

b) Given the following account balance, assume that there are two transactions – *Tran1* and *Tran2* that will be executed concurrently.

Account balance = 1000

Transaction	Operation
Tran1	Withdraw 500 from the account
Tran2	Deposit 800 into the account

Continued...

i) Explain how lost update may happen using concurrent execution schedule.

[2 marks]

ii) Explain how uncommitted data may happen using concurrent execution schedule.

[2 marks]

c) Deadlock occurs when two transactions wait for each other to unlock data. Wait/die and wound/wait schemes were used to prevent deadlock. Given *TranA* and *TranB* which start at *time1* and *time2* respectively, where *time1* is earlier than *time2*, answer the following questions.

i) Assume that *TranB* is owning the data item that is requested by *TranA*. What will happen when wait/die scheme is used to avoid deadlock?

[1 mark]

ii) Assume that *TranB* is owning the data item that is requested by *TranA*. What will happen when wound/wait scheme is used to avoid deadlock?

[1 mark]

iii)Between wait/die scheme and wound/wait scheme, which one is better? Why?

[1 mark]

Question 4 [10 marks]

a) In multimedia database, there are three retrieval approaches, namely attribute-based retrieval (ABR), text-based retrieval (TBR) and content-based retrieval (CBR). Describe TBR and explain TWO challenges of this approach.

[3 marks]

b) Assume that you are working in an animation and movie production company. You are asked to create a table for the *MOVIE* relation in your company's database and your table must be able to support ABR. The table also contains the video file. What are the FOUR relevant attributes and suitable data types that you can include in the table?

[2 marks]

Continued...

c) One of categories of NoSQL database systems is document-based systems. In document-based systems, collections of similar documents are stored. How does individual documents resemble XML documents? Give one example of document in XML or JSON format.

[3 marks]

d) Explain FOUR differences between data administrator and database administrator.

[2 marks]

End of page.